

12-05-00

#2

04CO
03-13-01
- OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/730,525
DATE: 03/19/2001
TIME: 13:26:22

Input Set : A:\09730525.txt
Output Set: N:\CRF3\03192001\I730525.raw

ENTERED

SEQUENCE LISTING

- 3 (1) GENERAL INFORMATION:
- 5 (i) APPLICANT: Ward, Eric
- 6 Volrath, Sandra
- 7 Johnson, Marie
- 8 Potter, Sharon
- 10 (ii) TITLE OF INVENTION: Herbicide Tolerant Protox Genes
- 11 Produced by DNA Shuffling
- 13 (iii) NUMBER OF SEQUENCES: 37
- 15 (iv) CORRESPONDENCE ADDRESS:
- 16 (A) ADDRESSEE: Novartis Corporation
- 17 (B) STREET: 3054 Cornwallis Road
- 18 (C) CITY: Research Triangle Park
- 19 (D) STATE: NC
- 20 (E) COUNTRY: USA
- 21 (F) ZIP: 27709
- 23 (v) COMPUTER READABLE FORM:
- 24 (A) MEDIUM TYPE: Floppy disk
- 25 (B) COMPUTER: IBM PC compatible
- 26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- 27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 29 (vi) CURRENT APPLICATION DATA:
- C--> 30 (A) APPLICATION NUMBER: US/09/730,525
- C--> 31 (B) FILING DATE: 05-Dec-2000
- 32 (C) CLASSIFICATION:
- 62 (vii) PRIOR APPLICATION DATA:
- 35 (A) APPLICATION NUMBER: 09/102,419
- 36 (B) FILING DATE:
- 39 (A) APPLICATION NUMBER: US 09/050,603
- 40 (B) FILING DATE: 30-MAR-1998
- 43 (A) APPLICATION NUMBER: US 09/038,878
- 44 (B) FILING DATE: 11-MAR-1998
- 47 (A) APPLICATION NUMBER: US 08/808,931
- 48 (B) FILING DATE: 28-FEB-1997
- 51 (A) APPLICATION NUMBER: US 60/012,705
- 52 (B) FILING DATE: 28-FEB-1996
- 55 (A) APPLICATION NUMBER: US 60/013,612
- 56 (B) FILING DATE: 28-FEB-1996
- 59 (A) APPLICATION NUMBER: US 60/020,003
- 60 (B) FILING DATE: 21-JUN-1996
- 63 (A) APPLICATION NUMBER: US 08/472,028
- 64 (B) FILING DATE: 06-JUN-1995
- 66 (viii) ATTORNEY/AGENT INFORMATION:
- 67 (A) NAME: Meigs, J. Timothy
- 68 (B) REGISTRATION NUMBER: 38,241
- 69 (C) REFERENCE/DOCKET NUMBER: CGC 1847/CIP4
- 71 (ix) TELECOMMUNICATION INFORMATION:

RAW SEQUENCE LISTING DATE: 03/19/2001
 PATENT APPLICATION: US/09/730,525 TIME: 13:26:22

Input Set : A:\09730525.txt
 Output Set: N:\CRF3\03192001\I730525.raw

```

72      (A) TELEPHONE: (919) 541-8587 .
73      (B) TELEFAX: (919) 541-8689
75 (2) INFORMATION FOR SEQ ID NO: 1:
77      (i) SEQUENCE CHARACTERISTICS:
78          (A) LENGTH: 1719 base pairs
79          (B) TYPE: nucleic acid
80          (C) STRANDEDNESS: single
81          (D) TOPOLOGY: linear
83      (ii) MOLECULE TYPE: cDNA
85      (iii) HYPOTHETICAL: NO
87      (iv) ANTI-SENSE: NO
89      (vi) ORIGINAL SOURCE:
90          (A) ORGANISM: Arabidopsis thaliana
92      (vii) IMMEDIATE SOURCE:
93          (B) CLONE: pWDC-2 (NRRL B-21238)
95      (ix) FEATURE:
96          (A) NAME/KEY: CDS
97          (B) LOCATION: 31..1644
98          (D) OTHER INFORMATION: /product= "Arabidopsis protox-1"
100     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
102 TGACAAAATT CCGAATTCTC TGCGATTTCC ATG GAG TTA TCT CTT CTC CGT CCG      54
103                                     Met Glu Leu Ser Leu Leu Arg Pro
104                                     1               5
106 ACG ACT CAA TCG CTT CTT CCG TCG TTT TCG AAG CCC AAT CTC CGA TTA      102
107 Thr Thr Gln Ser Leu Leu Pro Ser Phe Ser Lys Pro Asn Leu Arg Leu
108      10               15               20
110 AAT GTT TAT AAG CCT CTT AGA CTC CGT TGT TCA GTG GCC GGT GGA CCA      150
111 Asn Val Tyr Lys Pro Leu Arg Leu Arg Cys Ser Val Ala Gly Gly Pro
112      25               30               35               40
114 ACC GTC GGA TCT TCA AAA ATC GAA GGC GGA GGC ACC ACC ATC ACG      198
115 Thr Val Gly Ser Ser Lys Ile Glu Gly Gly Gly Gly Thr Thr Ile Thr
116      45               50               55
118 ACG GAT TGT GTG ATT GTC GGC GGA GGT ATT AGT GGT CTT TGC ATC GCT      246
119 Thr Asp Cys Val Ile Val Gly Gly Gly Ile Ser Gly Leu Cys Ile Ala
120      60               65               70
122 CAG GCG CTT GCT ACT AAG CAT CCT GAT GCT GCT CCG AAT TTA ATT GTG      294
123 Gln Ala Leu Ala Thr Lys His Pro Asp Ala Ala Pro Asn Leu Ile Val
124      75               80               85
126 ACC GAG GCT AAG GAT CGT GTT GGA GGC AAC ATT ATC ACT CGT GAA GAG      342
127 Thr Glu Ala Lys Asp Arg Val Gly Gly Asn Ile Ile Thr Arg Glu Glu
128      90               95               100
130 AAT GGT TTT CTC TGG GAA GAA GGT CCC AAT AGT TTT CAA CCG TCT GAT      390
131 Asn Gly Phe Leu Trp Glu Glu Gly Pro Asn Ser Phe Gln Pro Ser Asp
132      105               110               115               120
134 CCT ATG CTC ACT ATG GTG GTA GAT AGT GGT TTG AAG GAT GAT TTG GTG      438
135 Pro Met Leu Thr Met Val Val Asp Ser Gly Leu Lys Asp Asp Leu Val
136      125               130               135
138 TTG GGA GAT CCT ACT GCG CCA AGG TTT GTG TTG TGG AAT GGG AAA TTG      486
139 Leu Gly Asp Pro Thr Ala Pro Arg Phe Val Leu Trp Asn Gly Lys Leu

```

RAW SEQUENCE LISTING

DATE: 03/19/2001

PATENT APPLICATION: US/09/730,525

TIME: 13:26:22

Input Set : A:\09730525.txt

Output Set: N:\CRF3\03192001\I730525.raw

140	140	145	150	
142	AGG CCG GTT CCA TCG AAG CTA ACA GAC TTA CCG TTC TTT GAT TTG ATG	534		
143	Arg Pro Val Pro Ser Lys Leu Thr Asp Leu Pro Phe Phe Asp Leu Met			
144	155	160	165	
146	AGT ATT GGT GGG AAG ATT AGA GCT GGT TTT GGT GCA CTT GGC ATT CGA	582		
147	Ser Ile Gly Gly Lys Ile Arg Ala Gly Phe Gly Ala Leu Gly Ile Arg			
148	170	175	180	
150	CCG TCA CCT CCA GGT CGT GAA GAA TCT GTG GAG GAG TTT GTA CGG CGT	630		
151	Pro Ser Pro Pro Gly Arg Glu Glu Ser Val Glu Glu Phe Val Arg Arg			
152	185	190	195	200
154	AAC CTC GGT GAT GAG GTT TTT GAG CGC CTG ATT GAA CCG TTT TGT TCA	678		
155	Asn Leu Gly Asp Glu Val Phe Glu Arg Leu Ile Glu Pro Phe Cys Ser			
156	205	210	215	
158	GGT GTT TAT GCT GGT GAT CCT TCA AAA CTG AGC ATG AAA GCA GCG TTT	726		
159	Gly Val Tyr Ala Gly Asp Pro Ser Lys Leu Ser Met Lys Ala Ala Phe			
160	220	225	230	
162	GGG AAG GTT TGG AAA CTA GAG CAA AAT GGT GGA AGC ATA ATA GGT GGT	774		
163	Gly Lys Val Trp Lys Leu Glu Gln Asn Gly Gly Ser Ile Ile Gly Gly			
164	235	240	245	
166	ACT TTT AAG GCA ATT CAG GAG AGG AAA AAC GCT CCC AAG GCA GAA CGA	822		
167	Thr Phe Lys Ala Ile Gln Glu Arg Lys Asn Ala Pro Lys Ala Glu Arg			
168	250	255	260	
170	GAC CCG CGC CTG CCA AAA CCA CAG GGC CAA ACA GTT GGT TCT TTC AGG	870		
171	Asp Pro Arg Leu Pro Lys Pro Gln Gly Gln Thr Val Gly Ser Phe Arg			
172	265	270	275	280
174	AAG GGA CTT CGA ATG TTG CCA GAA GCA ATA TCT GCA AGA TTA GGT AGC	918		
175	Lys Gly Leu Arg Met Leu Pro Glu Ala Ile Ser Ala Arg Leu Gly Ser			
176	285	290	295	
178	AAA GTT AAG TTG TCT TGG AAG CTC TCA GGT ATC ACT AAG CTG GAG AGC	966		
179	Lys Val Lys Leu Ser Trp Lys Leu Ser Gly Ile Thr Lys Leu Glu Ser			
180	300	305	310	
182	GGA GGA TAC AAC TTA ACA TAT GAG ACT CCA GAT GGT TTA GTT TCC GTG	1014		
183	Gly Gly Tyr Asn Leu Thr Tyr Glu Thr Pro Asp Gly Leu Val Ser Val			
184	315	320	325	
186	CAG AGC AAA AGT GTT GTA ATG ACG GTG CCA TCT CAT GTT GCA AGT GGT	1062		
187	Gln Ser Lys Ser Val Val Met Thr Val Pro Ser His Val Ala Ser Gly			
188	330	335	340	
190	CTC TTG CGC CCT CTT TCT GAA TCT GCT GCA AAT GCA CTC TCA AAA CTA	1110		
191	Leu Leu Arg Pro Leu Ser Glu Ser Ala Ala Asn Ala Leu Ser Lys Leu			
192	345	350	355	360
194	TAT TAC CCA CCA GTT GCA GCA GTA TCT ATC TCG TAC CCG AAA GAA GCA	1158		
195	Tyr Tyr Pro Pro Val Ala Ala Val Ser Ile Ser Tyr Pro Lys Glu Ala			
196	365	370	375	
198	ATC CGA ACA GAA TGT TTG ATA GAT GGT GAA CTA AAG GGT TTT GGG CAA	1206		
199	Ile Arg Thr Glu Cys Leu Ile Asp Gly Glu Leu Lys Gly Phe Gly Gln			
200	380	385	390	
202	TTG CAT CCA CGC ACG CAA GGA GTT GAA ACA TTA GGA ACT ATC TAC AGC	1254		
203	Leu His Pro Arg Thr Gln Gly Val Glu Thr Leu Gly Thr Ile Tyr Ser			
204	395	400	405	

RAW SEQUENCE LISTING

DATE: 03/19/2001

PATENT APPLICATION: US/09/730,525

TIME: 13:26:22

Input Set : A:\09730525.txt

Output Set: N:\CRF3\03192001\I730525.raw

```

206 TCC TCA CTC TTT CCA AAT CGC GCA CCG CCC GGA AGA ATT TTG CTG TTG      1302
207 Ser Ser Leu Phe Pro Asn Arg Ala Pro Pro Gly Arg Ile Leu Leu Leu
208      410      415      420
210 AAC TAC ATT GGC GGG TCT ACA AAC ACC GGA ATT CTG TCC AAG TCT GAA      1350
211 Asn Tyr Ile Gly Gly Ser Thr Asn Thr Gly Ile Leu Ser Lys Ser Glu
212 425      430      435      440
214 GGT GAG TTA GTG GAA GCA GTT GAC AGA GAT TTG AGG AAA ATG CTA ATT      1398
215 Gly Glu Leu Val Glu Ala Val Asp Arg Asp Leu Arg Lys Met Leu Ile
216      445      450      455
218 AAG CCT AAT TCG ACC GAT CCA CTT AAA TTA GGA GTT AGG GTA TGG CCT      1446
219 Lys Pro Asn Ser Thr Asp Pro Leu Lys Leu Gly Val Arg Val Trp Pro
220      460      465      470
222 CAA GCC ATT CCT CAG TTT CTA GTT GGT CAC TTT GAT ATC CTT GAC ACG      1494
223 Gln Ala Ile Pro Gln Phe Leu Val Gly His Phe Asp Ile Leu Asp Thr
224      475      480      485
226 GCT AAA TCA TCT CTA ACG TCT TCG GGC TAC GAA GGG CTA TTT TTG GGT      1542
227 Ala Lys Ser Ser Leu Thr Ser Ser Gly Tyr Glu Gly Leu Phe Leu Gly
228      490      495      500
230 GGC AAT TAC GTC GCT GGT GTA GCC TTA GGC CGG TGT GTA GAA GGC GCA      1590
231 Gly Asn Tyr Val Ala Gly Val Ala Leu Gly Arg Cys Val Glu Gly Ala
232 505      510      515      520
234 TAT GAA ACC GCG ATT GAG GTC AAC AAC TTC ATG TCA CGG TAC GCT TAC      1638
235 Tyr Glu Thr Ala Ile Glu Val Asn Asn Phe Met Ser Arg Tyr Ala Tyr
236      525      530      535
238 AAG TAAATGTAAA ACATTAAATC TCCAGCTTG CGTGAGTTTT ATTAAATATT      1691
239 Lys
242 TTGAGATATC CAAAAAAAAA AAAAAAAAAA      1719
245 (2) INFORMATION FOR SEQ ID NO: 2:
247     (i) SEQUENCE CHARACTERISTICS:
248         (A) LENGTH: 537 amino acids
249         (B) TYPE: amino acid
250         (D) TOPOLOGY: linear
252     (ii) MOLECULE TYPE: protein
254     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
256 Met Glu Leu Ser Leu Leu Arg Pro Thr Thr Gln Ser Leu Leu Pro Ser
257 1      5      10      15
259 Phe Ser Lys Pro Asn Leu Arg Leu Asn Val Tyr Lys Pro Leu Arg Leu
260      20      25      30
262 Arg Cys Ser Val Ala Gly Gly Pro Thr Val Gly Ser Ser Lys Ile Glu
263      35      40      45
265 Gly Gly Gly Gly Thr Thr Ile Thr Thr Asp Cys Val Ile Val Gly Gly
266      50      55      60
268 Gly Ile Ser Gly Leu Cys Ile Ala Gln Ala Leu Ala Thr Lys His Pro
269 65      70      75      80
271 Asp Ala Ala Pro Asn Leu Ile Val Thr Glu Ala Lys Asp Arg Val Gly
272      85      90      95
274 Gly Asn Ile Ile Thr Arg Glu Glu Asn Gly Phe Leu Trp Glu Glu Gly
275      100     105     110
277 Pro Asn Ser Phe Gln Pro Ser Asp Pro Met Leu Thr Met Val Val Asp

```

RAW SEQUENCE LISTING

DATE: 03/19/2001

PATENT APPLICATION: US/09/730,525

TIME: 13:26:22

Input Set : A:\09730525.txt

Output Set: N:\CRF3\03192001\I730525.raw

278	115	120	125
280	Ser Gly Leu Lys Asp Asp	Leu Val Leu Gly Asp	Pro Thr Ala Pro Arg
281	130	135	140
283	Phe Val Leu Trp Asn Gly	Lys Leu Arg Pro Val	Pro Ser Lys Leu Thr
284	145	150	155
286	Asp Leu Pro Phe Phe Asp	Leu Met Ser Ile Gly	Gly Lys Ile Arg Ala
287	165	170	175
289	Gly Phe Gly Ala Leu Gly	Ile Arg Pro Ser Pro	Pro Gly Arg Glu Glu
290	180	185	190
292	Ser Val Glu Glu Phe Val	Arg Arg Asn Leu Gly	Asp Glu Val Phe Glu
293	195	200	205
295	Arg Leu Ile Glu Pro Phe	Cys Ser Gly Val Tyr	Ala Gly Asp Pro Ser
296	210	215	220
298	Lys Leu Ser Met Lys Ala	Ala Phe Gly Lys Val	Trp Lys Leu Glu Gln
299	225	230	235
301	Asn Gly Gly Ser Ile Ile	Gly Gly Thr Phe Lys	Ala Ile Gln Glu Arg
302	245	250	255
304	Lys Asn Ala Pro Lys Ala	Glu Arg Asp Pro Arg	Leu Pro Lys Pro Gln
305	260	265	270
307	Gly Gln Thr Val Gly Ser	Phe Arg Lys Gly Leu	Arg Met Leu Pro Glu
308	275	280	285
310	Ala Ile Ser Ala Arg Leu	Gly Ser Lys Val Lys	Leu Ser Trp Lys Leu
311	290	295	300
313	Ser Gly Ile Thr Lys Leu	Glu Ser Gly Gly Tyr	Asn Leu Thr Tyr Glu
314	305	310	315
316	Thr Pro Asp Gly Leu Val	Ser Val Gln Ser Lys	Ser Val Val Met Thr
317	325	330	335
319	Val Pro Ser His Val Ala	Ser Gly Leu Leu Arg	Pro Leu Ser Glu Ser
320	340	345	350
322	Ala Ala Asn Ala Leu Ser	Lys Leu Tyr Tyr Pro	Pro Val Ala Ala Val
323	355	360	365
325	Ser Ile Ser Tyr Pro Lys	Glu Ala Ile Arg Thr	Glu Cys Leu Ile Asp
326	370	375	380
328	Gly Glu Leu Lys Gly Phe	Gly Gln Leu His Pro	Arg Thr Gln Gly Val
329	385	390	395
331	Glu Thr Leu Gly Thr Ile	Tyr Ser Ser Ser Leu	Phe Pro Asn Arg Ala
332	405	410	415
334	Pro Pro Gly Arg Ile Leu	Leu Leu Asn Tyr Ile	Gly Gly Ser Thr Asn
335	420	425	430
337	Thr Gly Ile Leu Ser Lys	Ser Glu Gly Glu Leu	Val Glu Ala Val Asp
338	435	440	445
340	Arg Asp Leu Arg Lys Met	Leu Ile Lys Pro Asn	Ser Thr Asp Pro Leu
341	450	455	460
343	Lys Leu Gly Val Arg Val	Trp Pro Gln Ala Ile	Pro Gln Phe Leu Val
344	465	470	475
346	Gly His Phe Asp Ile Leu	Asp Thr Ala Lys Ser	Ser Leu Thr Ser Ser
347	485	490	495
349	Gly Tyr Glu Gly Leu Phe	Leu Gly Gly Asn Tyr	Val Ala Gly Val Ala
350	500	505	510

VERIFICATION SUMMARY DATE: 03/19/2001
PATENT APPLICATION: US/09/730,525 TIME: 13:26:23

Input Set : A:\09730525.txt
Output Set: N:\CRF3\03192001\I730525.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:1066 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:7
L:2052 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=16
L:2262 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=18
L:2472 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=20
L:2660 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=22
L:2818 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=24